2015

Rejustifying Retributive Punishment on Utilitarian Grounds in Light of Neuroscientific Discoveries more than Philosophical Calisthenics!

Robert B. McCaleb

REJUSTIFYING RETRIBUTIVE PUNISHMENT ON UTILITARIAN GROUNDS IN LIGHT OF NEUROSCIENTIFIC DISCOVERIES MORE THAN PHILOSOPHICAL CALISTHENICS!

ROBERT B. MCCALEB*

I. INTRODUCTION

Recent discoveries in neuroscience show that ancient and widely-held popular beliefs about free will, decision making, and voluntary action are deeply flawed, and that these concepts are potentially reducible to discrete, observable chemical events

* Cleveland-Marshall College of Law, J.D. Candidate, 2015. This article received the 2013–2014 Howard L. Oleck Award for Distinguished Legal Writing by a Student (Best Overall Paper) from the Cleveland-Marshall College of Law faculty, as well as the 2013–2014 Best Associate Note Award from the Editors of the Cleveland State Law Review. The author would like to thank Dr. Dana Watts for everything, ever, Professor Donald Watts for always believing in the author, if not in Materialism, and Caitlin Hill for believing in both.

in the brain.[1] The classical[2] criminal law, however, presupposes the existence of practically unrestrained free will, and demands that it be exercised within certain boundaries and in (or not in) certain ways.[3] Accordingly, viewed broadly, classical criminal law and materialist neuroscience rely on philosophically irreconcilable explanations of the sources and causes of volitional behavior.

The Rotten Social Background defense ("RSB")—first described by Judge Bazelon in his partial dissent in *U.S. v. Alexander*[4]—provides an eminently clear lens for scrutinizing the seemingly inevitable convergence of classical criminal law and materialist neuroscience. This is because RSB is qualitatively different from classical defenses in several relevant ways.[5] First, RSB is uniquely interested in the causal relationship between environmental conditions and social behavior. It views the criminal actor as a sensory-input/response-output mechanism with an observable degree of probabilistic determinism governing her behavior.[6] This concept will be discussed more fully below, but presently, it is important to note that this is a point of view widely accepted in neuroscience, partially (though controversially) accepted in moral philosophy, and generally not accepted in the law. Second, RSB takes a very broad timeline into account. It connects environmental stimuli to an actor's later behavior in a way that other defenses do not.[7] Third, it relies on an interpretation of

---

[1] Hans H. Kornhuber & Luder Deecke, *Readiness for Movement—The Bereitschaftspotential Story*, CURRENT CONTENTS, no. 4, Jan. 22, 1990, at 14, *available at* http://www.garfield.library.upenn.edu/classics1990/A1990CH18100001.pdf (providing a concise and readable relation of the event by the main characters); Benjamin Libet, *Do We Have Free Will?*, 6 J. CONSCIOUSNESS STUDIES No. 8-9, 47-57 (1999) (describing the experiments, their methodology and results, and Libet's sense of their implications). These neuroscientific explanations of behavior proceed from a materialist foundation that attempts to formulate a comprehensive description of the mind and its operations without recourse to unverifiable abstractions. *See* Daniel C. Dennett, *The Self as a Responding—and Responsible—Artefact*, 1001 ANNALS OF THE N.Y. ACAD. SCI. 39 (2003); Paul Churchland, *Some Reductive Strategies in Cognitive Neurobiology*, XCV MIND 379, 279 (1986), *available at* http://w3.uniroma1.it/cordeschi/Articoli/churchland.htm.

[2] I will use the modifier "classical" consistently throughout this note to refer to the Criminal Law as it exists in the Anglo-American criminal system today. This Criminal Law I take to be explicitly moralistic in its proscriptions, compatibilist in its approach to free will, retributive in its justifications for punishment, and dualistic in its treatment of the mind body problem.

[3] H.L.A. HART, PUNISHMENT AND RESPONSIBILITY (Oxford Univ. Press 1968).

[4] United States v. Alexander, 471 F.2d 923, 957-65 (D.C. Cir. 1972) (Bazelon, J., dissenting).

[5] *Id.* at 958-60. Regarding the classical defenses like duress, self-defense, intoxication and insanity, the focus is often on the actor's free conduct leading up to the criminal event in question. This is sometimes so in a duress situation, where the defense can be undone if the actor was negligent in putting himself into a situation where duress was foreseeable. *See, e.g.*, Regina v. Sharp, 3 W.L.R. 1, 8-9, Q.B. 853, 860-61 (1987) (U.K.). Even in an insanity defense situation, the investigation focuses largely on the accused's capacities at and right around the moment of the crime. *See* Blake v. United States, 407 F.2d 908 (5th Cir. 1969).

[6] WAYNE R. LAFAVE, PRINCIPLES OF CRIMINAL LAW 165-71, 218-24 (2d ed. 2010); ARNOLD H. LOEWY, CRIMINAL LAW IN A NUTSHELL 129-32, 152-158 (5th ed. 2009).

[7] *Alexander*, 471 F.2d at 957-59; LAFAVE, *supra* note 6, at 130.

extremely abnormal responses to non-exceptional stimuli by otherwise essentially normal individuals (by dualistic, libertarian standards) as sub- or non-voluntary.[8] Thus, RSB pierces the semi-permeable boundary between legal "sanity" and legal "insanity."[9]

Each of these unique features is skeptically received by the classical criminal law mainly because of its fundamentally non-scientific, folk-psychological position on free will.[10] However, discoveries in contemporary neuroscience strongly support each of these features.[11] While chemical, causal, and deterministic brain events indisputably play a central role, RSB is virtually ignored as a frightening intrusion by materialism into the dualist sanctuary of the law. Yet RSB, as an indicator species for the health of the criminal law's philosophic ecosystem, cannot be discounted out of hand any longer. In fact, the reconciliation of law and materialism is both practically and morally obligatory. This note proposes that such a reconciliation might be best accomplished by rebuilding the otherwise discredited retributive justifications of punishment on essentially utilitarian foundations, an idea discussed extensively below at section C of the Discussion.

Merely re-explaining the theoretical underpinnings of retributive punishment without proposing practical changes would amount to a demonstration of linguistic calisthenics. Therefore, in conjunction with a theoretical reappraisal, there must also be several practical modifications to traditional retributive "punishment" that may serve to maximize the utility of a re-justified utilitarian retributivism.[12] However, these expansions are beyond the scope of this note. Rather, the main practical focus will be on crafting a model statute—in the style of the American Legal Institute's Model Penal Code—that brings the aery language of Bazelon's dissent into a more concrete form.

The second part will begin by examining the content and reasoning behind Bazelon's dissent. Next, it will familiarize the reader with the most important and relevant brain structures, their broadly significant functions, and their relation to volitional behavior. Further, the second part will lay out the essential theoretical positions under which the analysis in the third part will proceed. The third part will

---

[8] Joshua Greene & Jonathan Cohen, *For the Law, Neuroscience Changes Nothing and Everything*, 359 PHILOS. TRANSACTIONS ROYAL SOC'Y BIO. SCI. 1776 (2004).

[9] LAFAVE, *supra* note 6, at 130-35.

[10] Emad H. Atiq, *How Folk Beliefs About Free Will Influence Sentencing: A New Target for the Neuro-Determinist Critics of Criminal Law*, 16 NEW CRIM. L. REV. 449, 474-92 (2013) (describing the "commonsense 'folk' notion of freedom" as one that is strongly opposed to determinism and founded on the presumption of total human volition and responsibility for actions, regardless of external— or internal— influences on the actor).

[11] *See generally* MICHAEL MOORE, PLACING BLAME: A GENERAL THEORY OF THE CRIMINAL LAW (1997).

[12] For instance: the transformation of prisons into inpatient clinics for the treatment of the neuro-physically disadvantaged; the black-letter formulation of a revised RSB defense that permits mitigation *without* devaluing responsibility beyond palatable lower limits; and (3) the progressive opening of the law to utilitarian ideas with retributivist moral foundations. Operating in concert, such changes would have profoundly beneficial effects by decreasing recidivism, increasing the respect of the under-privileged (i.e. the poor) and de-privileged (i.e. the incarcerated) for the rule of law, and minimizing economic waste created by useless repeat retributive incarceration of individuals with especially undermined freedom of will.

begin with an assessment of RSB's reception (or, indeed, rejection) in the decades since *Alexander*. It will then consider several reasons for its manifestly cool reception. The third section will then set forth the causes of the irreconcilability of classical criminal law and materialism. Finally, the third section will lay out an experimental approach to the problem of punishment in a world without unrestrained free will in the form of a model statute.

II. BACKGROUND

This part begins with a short history of the development of Rotten Social Background. Next, it describes several exemplary brain-areas and explains their roles in volitional behavior and decision-making. Finally, it summarizes the positions of the main camps into which advocates for radical reform of the criminal law have divided.

A. A Brief History of Rotten Social Background

In 1972, Judge David Bazelon introduced the Rotten Social Background defense to the American legal world in his dissent in *United States v. Alexander*.[13] In *Alexander*, the defendant, a young black man, shot and killed a white man who hassled him in a restaurant and called him a "black bastard."[14] The defendant grew up in an impoverished, single-parent home in the unstable Watts neighborhood of Los Angeles.[15] Throughout his early life he was tormented by racism, abuse, and little meaningful opportunity for inclusion in ordinary society.[16] At trial, his defense counsel proposed to describe the nature and effects of his upbringing to the jury in support of mitigation.[17] The trial judge rejected the proposal categorically.[18] In his pained and disquieting dissent, Judge Bazelon concluded that the jury should have

---

[13] United States v. Alexander, 471 F.2d 923, 957-65 (D.C. Cir. 1972) (Bazelon, J., dissenting).

[14] *Id.* at 957.

[15] *Id.*

[16] *Id.*

[17] *Id.* at 958-59.

[18] *Id.* In a perfect example of the typical attitude taken toward RSB by the judiciary at large, the district judge rejected the proposal in the following exchange:

> THE COURT: . . . I will tell them it is not in any way a question of his rotten social background.
> DEFENSE COUNSEL: I object.
> THE COURT: You may.
> COUNSEL: May I state my reasons?
> THE COURT: You may.
> COUNSEL: I was talking in terms of the cause of his condition.
> THE COURT: No, you weren't sir. You were appealing in the most direct way to something that I am going to keep out of the courtroom, if I stay a Judge. I am not going to permit it to come in here.

Id. at 959 n.101.

been allowed to hear about the defendant's terrible early life.[19] Judge Bazelon set out four mutually exclusive, almost-equally problematic outcomes if RSB were considered:

> We can impose narrow and admittedly illogical limitations on the responsibility defense to insure that a defendant like Murdock will not be acquitted on the theory that he lacked responsibility. [Or,] [i]f we remove the practical impediments to Murdock's defense and he is, in fact, acquitted for lack of responsibility, he could be released from custody in spite of his apparent dangerousness . . . [Or,] we can strive to find a vaguely therapeutic purpose for hospitalization . . . . Finally, if there are no known or foreseeable techniques for "curing" someone like Murdock . . . we [could] confine him in exclusive reliance on a prediction of dangerousness. That confinement would be nothing more or less than unadorned preventive detention.[20]

From among these four distasteful choices, Judge Bazelon chose the second, mainly on compassionate grounds and based on his conception of fundamental justice and injustice.[21] He was not blind to the frightening consequences his decision, if adopted, could have.

B. Professor Delgado's Reformulation and Expansion

Thirteen years after *Alexander*, Professor Richard Delgado argued for a similar defense based on "severe environmental deprivation."[22] Professor Delgado offered several compelling reasons for the adoption of such a defense, chiefly that appalling deprivation during upbringing has known and predictable effects on later decision-making and lifestyle.[23] Thus, like Judge Bazelon, Professor Delgado approached the problem of severely disadvantaged offenders from a sociological point-of-view and through compassionate appeals for reform.[24] However, where Judge Bazelon's reasoning was mainly jurisprudential, Professor Delgado's was all but revolutionary in character and aspiration.[25]

---

[19] *Id.* at 961-65. Judge Bazelon felt that though the defendant was not insane under the prevailing legal test, the extreme deprivation he suffered had produced behavioral propensities that could not be fitted tidily into traditional criminal beliefs about responsibility and desert.

[20] *Id.* at 962-64.

[21] *Id.*

[22] Richard Delgado, *"Rotten Social Background": Should the Criminal Law Recognize a Defense of Severe Environmental Deprivation?*, 3 L. & INEQ. 9 (1985) [hereinafter Delgado, *Rotten*].

[23] *Id.*

[24] *Id.* at 75-83; *see also* Richard Delgado, *The Wretched of the Earth*, 2 ALA. C.R. & C.L. L. REV. 1 (2011) [hereinafter Delgado, *Wretched*].

[25] *See, e.g.*, Delgado, *Rotten*, *supra* note 22, at 81-83. Delgado does not openly embrace the leftist, progressive point of view, though he describes its argument with evident sympathy. Thus,

> [I]n the Marxist and critical criminologist view, vengeful, self-seeking behavior is a normal response to an oppressive social system and can only be ended when that

III. PHILOSOPHICAL ARGUMENTS

This part sets out a novel philosophical argument in defense of a rejustified retributivism. It begins with a brief survey of the recent developments in neuroscience that tend to undermine, to a greater or lesser extent, the prevailing notions that underpin the criminal law—notions like free will, responsibility for conduct, blame for bad acts, and desert of punishment. Then, it examines some of the main philosophical responses to the problem of diminished free will as demonstrated by the explosive discoveries of neuroscience. Though the philosophical controversies are exceedingly complex, this note takes Jonathan Cohen and Joshua Greene's *For the Law, Neuroscience Changes Nothing and Everything* as a model position representing broadly the utilitarian-materialist position that stands against the classical libertarian-retributivist system currently justifying punishment. Finally, this part sets out the content and argumentation for a novel construction of retributivist punishment. The argument postulates, in substance, that though retributivism has been slowly discredited by neuroscience, it should be retained as the predominant justification for punishment because, when compared to its only legitimate philosophical rival—utilitarianism—it actually, and surprisingly, possesses *greater* utility.

A. Introduction to the Philosophical Argument

In a series of experiments conducted in the early 1980s, the neurologist Benjamin Libet asked participants to press a button when they saw a light turn on.[26] By measuring activity in each subject's brain, Libet discovered that the conscious act of pressing the button was preceded by an unconscious rise in the "readiness potential" of certain neurons, beginning 550 milliseconds before the act.[27] Astonishingly, the subjects only became aware of the intention to act 350 milliseconds after the rise in readiness potential but 200 milliseconds *before* the act occurred.[28] For 200

> system is radically restructured. The alienated worker or unemployed RSB person is shorn of self-respect, has no attachment to the broader community, and is a constant source of misdirected rebellion through crime. Punishment of persons for crimes they commit out of economic and political necessity is, in this view, senseless and inequitable.

Id. at 81. Most of Professor Delgado's subsequent scholarship has continued to focus on the sociopolitical aspects of rotten social background. *E.g.*, Delgado, *Wretched*, *supra* note 24; Richard Delgado, *The Myth of Upward Mobility*, 68 U. PITT. L. REV. 879 (2007) [hereinafter Delgado, *Myth*]; Richard Delgado, *Rodrigo's Portent: California and the Coming Neocolonial Order*, 87 WASH. U. L. REV. 1293 (2010); Richard Delgado, *Zero-Based Racial Policies: An Evaluation of Three Best-Case Arguments on Behalf of the Nonwhite Underclass*, 78 GEO. L.J. 1929 (1990). Likewise, the great majority of subsequent works by other authors inspired by Judge Bazelon and Professor Delgado have focused on the political and sociological side of the defense, and have espoused a more-or-less open political leftism. *E.g.*, Paul H. Robinson, *Are We Responsible for Who We Are?: The Challenge for Criminal Law Theory in the Defense of Coercive Indoctrination and "Rotten Social Background*," 2 ALA. C.R. & C.L. L. REV. 52 (2011); Erik Luna, *Spoiled Rotten Social Background*, 2 ALA. C.R. & C.L. L. REV. 23 (2011); Mythri Jayaraman, *Rotten Social Background Revisited*, 14 CAP. DEF. J. 327 (2002).

[26] *See* Libet, *supra* note 1, at 47.

[27] *Id.* at 50-51.

[28] *Id.* at 48-51.

milliseconds, therefore, the subjects' brains apparently began unconscious production of the act before they "chose" to initiate it.[29] In the course of these experiments, Libet also found that, although the unconscious brain may *initiate* voluntary acts, there is a short period of time during which the conscious brain can "veto" them.[30] Thus, Libet himself declined to proclaim the disproof of free will.[31] However, the implications of his work were repeatedly taken up and expanded upon by materialist philosophers and legal scholars.[32]

Over the last three decades, neuroscience has continuously produced evidence supporting some degree of physio-chemical determinism in human behavior.[33] Some of that evidence strongly suggests that the conscious brain's "veto" power can be limited, interfered with, or even eliminated entirely. This suggestion undermines prevailing retributivist justifications of punishment, which depend on the legitimacy of free will attributions.[34] Accordingly, it is necessary to make a reasoned, careful choice about how to punish in a world with less and less free will.

Assuming that we do not wish to abandon punishment entirely, we could transition from our current retributivist system based mainly on moral prerogatives to a (more-or-less untried) utilitarian system based mainly on efficiencies. Yet, while all punishment could conceivably be justified on utilitarian grounds, utilitarian justifications are subject to grave practical and ethical problems. In short, because of its pliable over-emphasis on efficiencies, utilitarianism is easily molded to justify almost any punishment.[35] But there is a way to retain retributive justifications while

---

[29] Note that the distinction between the "conscious" and "unconscious" brain essentially refers to the distinction between prefrontal cortex and the rest of the brain, respectively. The term "conscious" in this context typically denotes self-awareness and a capacity for moral and abstract reflection of the sort seemingly unique to humans. Thus, a cat is conscious in that it is aware of environmental stimuli, but not "conscious" in the sense of being capable of moral judgments. Cats are amoral. For concise and readable descriptions of any area of the brain discussed in this note, see generally RITA CARTER, MAPPING THE MIND (rev. ed. 2010) [hereinafter CARTER, MAPPING].

[30] Libet, *supra* note 1, at 52-54.

[31] *Id.* at 55-57.

[32] Materialism, or "physicalism," is the thesis that everything is physical— made of matter—including traditionally non-physical phenomena like the mind. If everything is physical, the argument goes, then everything is subject to the laws of physics and causation—from the formation of a star to the formation of a thought. *See* Daniel Stoljar, *Physicalism*, *in* THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta, ed.) (Feb. 13, 2001, rev. Sept. 9, 2009), http://plato.stanford.edu/entries/physicalism/ (last visited Nov. 2, 2014). For examples of materialists using Libet's experiments to prove their point, see, for example, Dennett, *supra* note 1 (arguing that evidence of preconscious decision-making undermines the plausibility of causally-effective conscious persons); Churchland, *supra* note 1 (using neuroscience to bolster his hardcore, almost fanatical materialist position that, like Dennett's, goes almost all the way to complete denial of causally effective personhood). The law itself remains mostly unaffected. *See infra* Part II.C.

[33] *See generally* MICHAEL S. GAZZINAGA, WHO'S IN CHARGE?: FREE WILL AND THE SCIENCE OF THE BRAIN (reprt. ed. 2012).

[34] *See infra* Part II.A.

[35] *See infra* Part III.A.

recognizing, and even incorporating, neuroscientifically-supported circumscription of free will: it is to re-justify retributivism on *utilitarian* grounds. This is more than a philosophical parlor trick—such a re-conception would allow greater inclusivity for novel defenses, buttress the legitimacy of the criminal law against future neuroscientifically-derived materialist attacks, and ensure that punishment is predictably and accurately dispensed. This approach follows from the postulate that it is better for society to treat people as if they have unencumbered free will—even if they might not—than to treat them like chemical mechanisms controlled by the laws of physics—even if they might be. It amounts to saying that retributivism has more utility than utilitarianism.

B. Terminology of the Criminal Law and Philosophical Concepts Underlying It

It is black letter law that (in almost every case) criminal liability requires both *mens rea* and *actus reus*.[36] Because the *actus reus* requirement is generally understood to entail (more-or-less) philosophically unrestrained free will, many scholars have described the criminal law as metaphysically "libertarian."[37] Determinism, the counter-point to libertarianism, basically holds that the identifiable causal chain behind every past and present event permits theoretically perfect prediction of future events, including complex human behaviors.[38] If perfect prediction is possible, because of the laws of cause and effect, then actors are actually heavily restricted in their "choices."[39] Accordingly, deterministsm—who are almost always also materialists—have serious doubts about the scope of free will and, by extension, retributivism.[40]

It is also generally agreed that there are two chief ways of justifying punishment – Kantian retributive justifications[41] and Benthamite utilitarian justifications.[42]

---

[36] The actus reus must be a "voluntary" act, and the actor must commit that voluntary act with the requisite mens rea. Except in very rare circumstances, the actor must at least act "negligently" with respect to a material element of the offense. *See* LAFAVE, *supra* note 6, at 165-71, 218-24; LOEWY, *supra* note 6, at 129-32, 152-158.

[37] Libertarianism as a metaphysical position is distinct from and unrelated to the political theory of the same name. *See* Peter Vallentyne, *Libertarianism*, *in* THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward L. Zalta ed.) (Sept. 5, 2002, rev. July 1, 2014), http://plato.stanford.edu/entries/libertarianism/ (last visited Nov. 2, 2014); *see generally* HART, *supra* note 3 (addressing throughout, and attempting to refute, determinist critiques of the criminal law).

[38] *See* Carl Hoefer, *Causal Determinism*, *in* THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward L. Zalta ed.) (Jan. 23, 2009, rev. Jan. 21, 2010), http://plato.stanford.edu/entries/determinism-causal/ (last visited Nov. 2, 2014).

[39] That is, a "choice" proceeds from a cause, and is merely that cause's inevitable effect.

[40] Andrew E. Lelling, *Eliminative Materialism, Neuroscience and the Criminal Law*, 141 U. PA. L. REV. 1471 (1993) (attempting to describe a theoretical utilitarian legal system in a world without free will). *But see generally* BRAND BLANSHARD, REASON AND BELIEF (Yale Univ. Press 1974) (an idealist, as opposed to a materialist, attempting to reconcile determinism and responsibility, but only incidentally with an eye to the criminal law).

[41] *See* IMMANUEL KANT, GROUNDWORK OF THE METAPHYSICS OF MORALS 62-66 (H.J. Paton, trans., Wilder Publ'ns 2008) (establishing the fundamental parts of his famous "categorical imperative" of moral duties).

[42] *See, e.g.*, JEREMY BENTHAM, THE PRINCIPLES OF MORALS AND LEGISLATION (Prometheus Books 1988) (one of Bentham's many works arguing from the assumption that "the good" is

Retributive justifications broadly hold that punishment is justified because those who break the law *deserve* to be punished.[43] Utilitarian justifications, on the other hand, broadly hold that punishment is justified whenever it has beneficial effects on society, and not necessarily when it is "fair" or only when the lawbreaker "deserves" it.[44] Punish in American criminal law, according to leading criminal law scholars, has been justified mostly on retributive grounds.[45] However, as neuroscience matures, the fundamental assumptions underpinning classical retributive justice have been more and more tightly circumscribed.[46] Thus, the philosophical tide has been shifting towards utilitarian theories, as these do not require unrestrained (or any) free will because they do not fundamentally require that an actor deserve punishment.[47]

---

that which increases the happiness of the maximum number of people); *see also* JEREMY BENTHAM, PANOPTICON; OR, THE INSPECTION HOUSE (Kessinger Publ'g 2010) (proposing a theory that eventually became, in modified form, the penal norm); MICHEL FOUCAULT, DISCIPLINE AND PUNISH: THE BIRTH OF THE PRISON 195-230 (Vintage Books 1977) (criticizing Bentham's penal theories and tracing their history and development in the West); Stephen J. Morse, *Excusing and the New Excuse Defenses: A Legal and Conceptual Review*, 23 CRIME & JUST. 329, 345 (1998) ("There is no consensually accepted meaning of determinism, but a typical understanding is that the laws of the universe and antecedent events together determine all future events. Many people assume that this is true, at least at levels higher than the explanation of subatomic particles, and it is certainly the background assumption of many working scientists.").

[43] *See* MOORE, *supra* note 11, at 83-94 (describing the various theories of punishment).

[44] *See, e.g.*, Greene & Cohen, *supra* note 8, at 1776-77 (explaining and defending prevailing conceptions of utilitarianism).

[45] *See generally* Anders Kaye, *The Secret Politics of the Compatibilist Criminal Law*, 55 U. KAN. L. REV. 365 (2007).

[46] *See* Anders Kaye, *Powerful Particulars: The Real Reason the Behavioral Sciences Threaten Criminal Responsibility*, 37 FLA. ST. U. L. REV. 539, 546 (2010).

> It is true, of course, that nothing in the behavioral sciences is sufficient to establish the truth of determinism outright. Behavioral science findings do not come close to mapping a complete universe of causal explanations for human acts, and nothing in the research—standing on its own—makes it obvious that such a map is inevitable. But we might naturally take the growing body of behavioral science findings to make an important contribution to the case for determinism in human action. Our lives are full of causal stories about the events that make up our days . . . . [T]he behavioral sciences seem to tell us, the machinery of causation is everywhere to be found. Thus, while nothing in the behavioral sciences 'proves' the truth of determinism, behavioral science findings may nevertheless contribute to the case for determinism.

Id.

[47] *See* Anders Kaye, *Resurrecting the Causal Theory of the Excuses*, 83 NEB. L. REV. 1116, 1124-25 (2005) (arguing against certain strains of retributivism that "hold[] that there can be no responsibility without control: 'if an action results from a deterministic causal process that traces back to factors beyond the control of the agent, he is not morally responsible for the action'; if an actor is to be blamed for his conduct, he 'should be ultimately responsible . . . "ultimately" in the sense that nothing for which [he was] not responsible should be the source of [his] [conduct]'; 'If agents' acts are caused by factors for which they are not responsible, then how can they be morally responsible for acting as a result of those factors?' In short, conduct caused by forces or circumstances the actor does not control is not sufficiently 'free' to justify attributions of moral responsibility.").

They require only that some valuable societal aim be accomplished by punishing him. What are some of the neuroscientific discoveries supporting this shift?

C. An Introduction to Two Brain Areas Whose Functions Undermine the Legitimacy of Free Will Attributions—The Amygdala and the Anterior Cingulate Gyrus

Though neuroscientific discoveries have expanded our understanding of every area of the brain, two areas are of particular illustrative value in this context—the amygdala and the anterior cingulate gyrus.[48] These areas are both in the "unconscious" areas of the brain, yet they exert great, frequently overriding influence over the conscious, decision-making cortex.

The amygdala is a matched pair of almond-shaped neuronal knots embedded deep in the temporal lobes of the brain.[49] Its location between the ancient, primitive structures of the midbrain and the relatively modern prefrontal cortex—which is the seat of higher brain functions like cognition and abstract reasoning—gives it an important, unconscious mediating function.[50] It is believed to play an extremely important role in regulating emotional responses produced by the midbrain to environmental stimuli perceived by the cortex.[51] It responds more quickly to environmental stimuli than the conscious brain, and its signals often overwhelm the slower, more deliberate cortex—especially in high stress, fight-or-flight situations. Thus, the amygdala's influences on anger, aggressiveness, and spontaneous violence appear to be especially substantial.[52]

Nearby is the anterior cingulate gyrus. This area is a bow-shaped region of the forebrain, located just under the wrinkled surface of the conscious prefrontal cortex.[53] Though its chief functions are to regulate autonomic bodily functions like heart rate and intestinal contractions, it also appears to have a central role in

---

[48] Though there are other brain areas directly related to behavior and willed action, these two suffice to demonstrate the increasing understanding of the role *biology* plays in events traditionally explained by morality and philosophy.

[49] CARTER, MAPPING, *supra* note 29, at 15.

[50] Thus, stimuli that cause automatic fear responses—like angry human faces, dangerous animals, or visible weapons—are processed by the prefrontal cortex and shunted to the motor cortex, with unconscious, nearly concurrent regulatory input from the fight-or-flight limbic system (especially the amygdala) ultimately mediating the conscious response. *See id.* at 93-97.

[51] *See id.* Thirty years of research has confirmed this view repeatedly. *See* Larry Swanson & Gorcia D. Petrovich, *What is the Amygdala?: A Comparative Approach*, TRENDS IN NEUROSCIENCE, Aug. 1998, 323, 323-24; Elisabeth A. Murray, *The Amygdala, Reward, and Emotion*, 11 TRENDS IN COGNITIVE SCI. 489, 491-97 (2007).

[52] Especially when the amygdala is damaged or malformed, or when the neural connections from the amygdala to the prefrontal cortex are malfunctioning or imbalanced. *See, e.g.*, Justin S. Feinstein et al., *The Human Amygdala and the Induction and Experience of Fear*, 21 CURRENT BIOLOGY 34, 34 (2011); Justin S. Feinstein et al., *Fear and Panic in Humans with Bilateral Amygdala Damage*, 16 NATURE NEUROSCIENCE 270, 270 (2013); Daniel Kennedy et al., *Personal Space Regulation by the Human Amygdala*, 12 NATURE NEUROSCIENCE 1226, 1226 (2009).

[53] RITA CARTER, THE HUMAN BRAIN 121-26 (2009) [hereinafter CARTER, BRAIN].

initiating higher cognitive events like decision-making and impulse control.[54] Together, the anterior cingulate gyrus and the amygdala are central to human behavioral production and control. Both areas, however, are distinct from, but connected to, the conscious areas of the brain.[55] As noted above, these are assuredly not the only brain areas linked to unconscious behavior creation and control. Rather, they exemplify neuroscientifically-supported discoveries tending to show that most human behavior is caused and governed by chemical events in the body.

D. The Current State of the Law in this Context

Criminal case law has been extremely hesitant to entertain materialist, neuroscientifically-supported arguments. An early example comes from *Pope v. United States*,[56] wherein the United States Supreme Court declined to allow psychiatric testimony that the defendant suffered from an uncontrollable impulse to consume alcohol.[57] Justice Black, concurring with the majority, produced an excellent example of judicial terror at the inclusion of neuroscientific background evidence.[58] He wrote that by allowing psychiatric evidence of the kind proffered "the Court would be forced to hold the States powerless to punish any conduct that could be shown to result from a 'compulsion.'"[59] Even in the late 1960s, then, well before the flourishing of neuroscience, there was recognition that the slightest opening for neuroscientific determinism would unleash a torrent. Three decades later, in *Buchanan v. Angelone*, the United States Supreme Court restated this general trepidation when it upheld a state court's refusal to order jury instructions on the effects of defendants' backgrounds on their decision-making abilities.[60] Likewise,

---

[54] *See, e.g.*, George Bush et al., *Cognitive and Emotional Influences in Anterior Cingulate Cortex*, 4 TRENDS IN COGNITIVE SCI. 215, 215 (2000). The anterior cingulate gyrus' role in impulse control has stimulated some scientific interest in its role in criminality and recidivism. *E.g.*, Eyal Aharoni et al., *Neuroprediction of Future Rearrest*, 110 PNAS 6223, 6223 (2013). Francis Crick, the famed co-discoverer of DNA, describes the anterior cingulate gyrus as a likely candidate for the source of willed action. *See* FRANCIS CRICK, THE ASTONISHING HYPOTHESIS 129-32 (1995).

[55] *See* CARTER, BRAIN *supra* note 53, at 121-26.

[56] 392 U.S. 651, 1968 LEXIS 1161 (1968).

[57] Powell v. State of Texas, 392 U.S. 514, 536-37 (1968) (preferring, however, to resolve the issues on constitutional grounds).

[58] *Id.* at 535 (Black, J., concurring).

[59] *Id.* at 543; *see also* United States v. Alexander, 471 F.2d 923, 957-65 (D.C. Cir. 1973) (Bazelon, C.J., dissenting). In another perfect example of the typical attitude taken by the judiciary at large toward non-insanity mental background evidence, the district judge (quoted by Bazelon, and cited above) rejected a proposal to instruct the jury on the determinative effects of the defendant's mental and psychiatric background.

[60] Buchanan v. Angelone, 522 U.S. 269, 273-74, 279 (1998); *cf.* Sexton v. State, 997 So. 2d 1073, 1083-84 (Fla. 2008) (reasoning that PET scan evidence of limbic system deficiencies suggesting possible interference with intent formation was admissible, but intimating that its probative worth was extremely limited). The limbic system contains the amygdala, described above.

courts continue not to permit evidence of genetic predispositions to violence[61] or chromosomal abnormalities increasing reactive aggression,[62] and even question the validity of inquiries into *any* mental abnormalities not rising to the level of legal insanity.[63] As far as the law is concerned, a person "is either sane or insane; there is no sliding scale . . . ."[64]

Thus, the law maintains a deep mistrust towards materialism, even as neuroscience matures.[65] Because our libertarian-retributivist system relies on the legitimacy of free will attributions, the least assault on the validity of the concept is taken as a brazen attack on the entire edifice of the law.[66] The risks of utilitarianism—long considered the only plausible alternative to classical retributivism—make this judicial terror understandable. But this reactionary attitude is not warranted, as we shall see below.

E. The Utilitarians and the Retributivists

Part III continues with a survey of recent scholarship on the problem neuroscience poses, focusing on Joshua Greene and Jonathan Cohen's *For the Law,*

---

[61] Matthew L. Baum, *The Monoamine Oxidase A (MAOA) Genetic Predisposition to Impulsive Violence: Is It Relevant to Criminal Trials* (2011), http://www.pc.rhul.ac.uk/sites/rheg/wp-content/uploads/2011/12/genetic-italy-case.pdf (examining changing trends in Europe, and answering in the affirmative the question posed in its title).

[62] *See* United States v. Eff, 524 F.3d 712, 718-20 (5th Cir. 2008) (holding expert psychiatric/neuroscientific testimony irrelevant if not rising to level of insanity defense). *But see* Cornwell v. Bradshaw, 559 F.3d 398, 418-20 (6th Cir. 2009) (Moore, J., dissenting) (arguing that failure to allow *any* expert testimony on XXY syndrome was an abuse of discretion).

[63] *See* Stamper v. Commonwealth, 324 S.E.2d 682, 688 (Va. 1985) (holding that "evidence of a criminal defendant's mental state at the time of the offense is, in the absence of an insanity defense, irrelevant to the issue of guilt."); Peeples v. Commonwealth, 519 S.E.2d 382, 384-85 (Va. Ct. App. 1999) (reaffirming the "unsuitability of psychiatry for determining criminal responsibility in the absence of an insanity defense" because of its dynamic, uncertain theories).

[64] *Peeples*, 519 S.E.2d at 384.

[65] As one scholar put it,

> [t]he criminal law generally rejects excuses based on causal accounts of criminality. For example, we know from personal experience and scientific study that much of our basic temperament derives from—that is, is caused by—genetic inheritance and so provides a classic instance of unchosen influence upon action. If the bad-tempered person feels a much greater temptation to violence in aggravating situations than do others, the difference stems from genetics, not free choice. Nevertheless, the ill-tempered person cannot base a provocation claim on personality type, nor can the absent-minded claim exemption from reckless or negligent conduct. The result is no different if the offender presents evidence that his basic disposition toward doing wrong resulted from unchosen environmental influences.

Samuel H. Pillsbury, *The Meaning of Deserved Punishment: An Essay on Choice, Character, and Responsibility*, 67 IND. L.J. 719, 729 (1992).

[66] Again, take the position of the district judge in *Alexander* as a typical example.

Neuroscience Changes Nothing and Everything.[67] Part III then examines the materialist-utilitarian position, as embodied by their article. Part III concludes by showing that neuroscience and retributivism can be reconciled by replacing the moral philosophy underlying retributivism with utilitarian reasoning.

1. The Main Thrust of the Materialist-Utilitarian Argument

> Was it him, or was it his circumstances? Was it him, or was it his brain? [W]hat most people do not understand, despite the fact that naturalistic philosophers and scientists have been saying it for centuries, is that there is no 'him' independent of these other things. (Or, to be a bit more accommodating to the supernaturally inclined, there is no 'him' independent of these things that shows any sign of affecting anything in the physical world, including his behaviour.)
>
> —Joshua Greene and Jonathan Cohen, *For the Law, Neuroscience Changes Nothing and Everything*, 359 Phil. Transactions Royal Soc'y Biological Sci. 1776 (2004).

Joshua Greene and Jonathan Cohen have produced one of the most cogent and farsighted defenses of neuroscientifically-supported utilitarianism.[68] Greene and Cohen take an expansive view of materialism. In their view, "there is not a shred of scientific evidence to support the existence of *causally effective* processes in the mind or brain that violate the laws of physics."[69] Thus, any cogent discussion of free will, they argue, must proceed from the assumption that free will does not exist in a way the law traditionally requires.[70] First, they demolish the philosophical foundations of an increasingly "panicky libertarianism," and proceed to dismiss as metaphysically incoherent the "compatibilist" position that free will and determinism are somehow reconcilable.[71] They daringly claim that

> [a]rguments are nice, but physical demonstrations are far more compelling. What neuroscience does, and will continue to do at an accelerated pace, is elucidate the 'when', 'where' and 'how' of the mechanical processes that cause behavior. It is one thing to deny that human decision-making is purely mechanical when your opponent offers only a general, philosophical argument. It is quite another to hold your ground when your opponent can make detailed predictions about how these mechanical processes work, complete with images of the brain structures involved and equations that describe their function.[72]

---

[67] Greene & Cohen, *supra* note 8.

[68] *Id.* For the sake of argument, this note assumes their position is broadly representative, and critiques its assumptions as if they were generally held by utilitarians.

[69] *Id.* at 1777. This position is essentially an argument from Libet's findings.

[70] *Id.* at 1777-78.

[71] *Id.* at 1780.

[72] *Id.* at 1781.

Then, Greene and Cohen argue that utilitarianism *alone* is capable of justifying punishment in a world without free will while simultaneously and embracing neuroscientifically-supported materialism.[73] They argue that because utilitarianism does not distinguish between the metaphysically responsible actor (who they call a "systematic illusion") and the irresponsible one, it is able to dispense punishment with much greater accuracy and effectiveness than retributivism.[74] Thus, in essence, their position is that utilitarianism—or, as they call it, "consequentialism"—is superior to retributivism because utilitarianism is uniquely capable of differentiating between who should and who should not (or need not) be punished.[75] Those whose punishment advances a societal aim should be punished; those whose punishment accomplishes nothing—because they are philosophically irresponsible by virtue of neuroscientific determinism—should not be.

Greene and Cohen spend a large portion of their article responding to critics who argue that utilitarianism makes extreme over-punishment troublingly easy to justify.[76] They contend that extreme over-punishment is implausible for two related reasons.[77] First, they argue that useful but extreme over-punishment is unlikely because its utility will almost always be outweighed by the attendant disruptions and anxieties ostensibly useful extreme punishments would create.[78] They then argue that utilitarian justifications, if widely accepted, would actually have the opposite effect.[79] That is, rather than extreme over-punishment, utilitarian justifications would allow legal decision-makers to make *more* compassionate decisions about desert and responsibility than those possible under a classical retributivist system. As neuroscientific materialism becomes widely accepted by the lay populace, legal decision-makers (lay and trained alike) would recognize that there is little direct utility in punishing criminals who cannot help committing crimes if their behavior is predetermined.[80] They conclude their article optimistically, saying that

> [f]ree will as we ordinarily understand it is an illusion generated by our cognitive architecture . . . . At this time, the law deals firmly but mercifully with individuals whose behaviour is obviously the product of forces that are ultimately beyond their control. Some day, the law may treat all convicted criminals this way. That is, humanely.[81]

It is not clear how this system would self-correct. But more troubling, though Greene and Cohen obviously disagree, is that utilitarianism *is* subject to grave and

---

[73] *Id.*

[74] *Id.* at 1778-80, 1783.

[75] *Id.*

[76] *Id.* at 1783. Greene and Cohen's example of possible extreme over-punishment involves capital punishment for parking violations. No doubt this would dramatically lower the number of parking violations. But at what cost?

[77] *Id.*

[78] *Id.*

[79] Greene & Cohen, *supra* note 8, at 1783.

[80] *Id.* at 1784.

[81] *Id.*

incurable doubts—just not the straw-man doubts they raise. As noted above, critics of utilitarianism point out that under a utilitarian system the government could rationally mandate, say, street-side executions for parking violations or other draconian penalties purportedly resulting in a net increase in safety and efficiency. Greene and Cohen optimistically argue that such schemes would not even be proposed, much less accepted, but their analysis is probably over-confident on this point. First, governments establish monstrous but ostensibly "useful" penalties relatively easily and often.[82] Second, extreme penalties are easy to dismiss as improbable straw men, but punishments predicated entirely on efficiency need not be *extreme* to be *unjust*.

For instance, under a utilitarian system the government could justify punishing people for relatively mild crimes like tax fraud *even if they are innocent* on a purely deterrent basis.[83] Or *not* punish people for some offenses as long as their punishment could be convincingly faked. Even more disturbing examples come to mind, all of which are far more plausible than Greene and Cohen's capital-punishment hypothetical. For instance: performing involuntary medical testing on death-row or life-sentence inmates; forced sterilization of sex criminals; even preventive detention of people with certain genotypes, mental disorders, or communicable diseases.

Utilitarianism allows the efficient and cost-effective elimination of destabilizing, troublesome, and undesirable elements in society, and is arguably superior to retributivism in this one regard. The problem is that it contains no static, internal definitions of labels like "destabilizing," "troublesome," or "uncomfortable." From one moment to the next—or one ruling class to the next—there is no semi-permanent basis for agreeing generally on *who* is "destabilizing" (and therefore subject to punishment for the good of the rest of the populace). One day your enemy is a destabilizing force. Tomorrow you could be.

Thus, utilitarian justifications are untrustworthy and dangerous in part because of their greater susceptibility to subjective rationalization than retributive justifications, but also in part because of their more cavalier attitude towards the cherished retributive folk-belief that "the punishment should fit the crime." It cannot be overemphasized that punishment need not be barbarous or sadistic to be unjust. But

---

[82] For instance, the still-unremedied and manifestly unfair crack-powder cocaine sentencing disparities are "useful" insofar as they permit extremely long detentions for making, selling, or using a dangerous and community-destroying drug (and, arguably, dealing effectively with surplus population). United States v. Blewett, 719 F.3d 482 (6th Cir. 2013). The detention of Japanese-American during World War II was "useful" insofar as the program quelled public fears and ostensibly (perhaps actually) prevented fifth-column attacks. Korematsu v. United States, 323 U.S. 214 (1944), *reh'g denied,* 324 U.S. 885 (1945). The water-boarding, enhanced "interrogation", force-feeding, extraordinary rendition, and indefinite detention of suspected terrorists are all useful in obtaining military intelligence and advanced warning of future attacks. Padilla v. Hanft, 423 F.3d 386 (4th Cir. 2005). Though utilitarian justifications are the exception, rather than the norm, they are—like the examples above—almost always objectionable despite their incontestable utility. Note, also, that "utility" *is* almost always defined exclusively by the ruling class. *See* Delgado, *Wretched*, *supra* note 24, at 9.

[83] Judge Bazelon was especially troubled by the possibility of preventive detention defended on determinist grounds. *See* United States v. Alexander, 471 F.2d 923, 964 (D.C. Cir. 1973) (Bazelon, C.J., dissenting).

that leaves us to choose between discredited utilitarianism and discredited retributivism. Or does it?

2. The Case for (Quiet) Change

Retributivism, even if empirically questionable, and at times inefficient, makes desirable demands of individual actors. It insists that we use our free will in a way that comports with the law, and punishes us when we do not—on the theory that when we disobey the law it is because we are bad. Most people fulfill these demands, and fulfilling them has positive results. For instance, retributivism, because of its emphasis on individual responsibility and objective moral judgments, allows for rational long-term planning, and creates predictability and order based on an assumption that people can and will make really free choices that comport with the law. The problem, of course, is that really free choices may not actually exist.[84] If real choices do not exist then retributive justifications founded upon the existence of real choices is *facially* indefensible. It would then appear that we are left to choose[85] between utilitarianism (and its attendant risks) and apparently discredited, empirically insupportable retributivism.

The better choice is an empirically false, but socially invaluable retributivism, because *retributivism has a higher utility than utilitarianism itself*. As shown above, retributivism is subject to philosophical doubts, but utilitarianism is subject to practical and ethical doubts. Retributivism is virtually incapable of manipulation except insofar as moral attitudes shift slowly over time, while utilitarianism is unpredictable and subject to the shifting needs of the ruling class or political majority.[86] Retributivism is subject only to the general folk conceptions (even if *false*) of human dignity, freedom, and responsibility.[87] In short, retributivism is based on possibly false beliefs that create a more desirable state of society than possible

[84] Or may exist in a way that makes the actor doing the choosing morally blameless, as an extension of Libet's results would suggest. *See supra* Part II.A-B. *See generally* Gideon Yaffe, *Libet and the Criminal Law's Voluntary Act Requirement*, *in* CONSCIOUS WILL AND RESPONSIBILITY: A TRIBUTE TO BENJAMIN LIBET (Walter Sinnott-Armstrong & Lynn Nadel eds., 2011).

[85] If the reader will pardon the expression.

[86] *See generally* Stephen Morse, *Mental Disorder and Criminal Law*, 101 J. CRIM. L. & CRIMINOLOGY 885 (2011) (arguing that neuroscience can demonstrate nothing about that world that would actually undermine the *moral* foundations of retributivism); Stephen J. Morse, *Excusing and the New Excuse Defenses: A Legal and Conceptual Review*, 23 CRIME & JUST. 329, 344-45 (1998) ("[M]ost attacks on the general justification of responsibility and blaming claim either that the truth of determinism renders the whole enterprise incoherent or that the criminal law is deeply inconsistent about the effect of determinism on responsibility, an inconsistency that cannot be satisfactorily resolved if determinism is true. Perhaps the most common response to such claims is to reduce responsibility and blaming to purely forward-looking, consequentially justified practices that depend on the truth of determinism for their efficacy. [But the] negative claims are unwarranted and that the usual response is itself unsatisfactory as a justification of our practices . . . .").

[87] *See generally* HART, *supra* note 3 (setting out and defending many traditional notions of responsibility, blame, and desert).

under utilitarianism, even if utilitarianism has more empirical support through its links to materialism and neuroscience.[88]

A re-justified retributivism scheme, defended on the grounds of its superior usefulness, would allow modifications to the criminal law that were thought dangerous precedents or unrestricted extensions to the general irresponsibility defenses. This is because a retributive system self-justified on the basis of its usefulness could openly embrace neuroscientifically-derived changes without fear, but *only* when they are useful in producing predictable results and not only when they are useful in maximizing the "greater good." Thus, Greene and Cohen's lay-perceptual shift could still occur (and have beneficial effects) because actors would be punished as if they had free will only when a free will attribution advanced the reasonable demands retributivism *per se* placed on that actor. Lay biases based on retributive folk psychology sometimes do result in injustices, but these would only be exacerbated if the criminal law did not have to worry about what is believed to be right or wrong and only about what is believed to be useful.[89]

At the end of the day, it is more useful to believe in right and wrong—to act as if right and wrong exist and one can know them and choose from between them, to insist that others do the same, and to punish them when they do not—than it is to admit that we are extremely complex, semi-enclosed, self-perpetuating chemical reactions with no causally-effective capability to avoid the laws of physics, and no verifiable free will.[90] The latter view may correspond more directly to the truth. But

---

[88] *See generally* THOMAS NAGEL, MIND AND COSMOS: WHY THE NEO-DARWINIAN CONCEPTION OF NATURE IS ALMOST CERTAINLY FALSE (2012) (arguing that materialism—which Nagel calls "neo-darwinism"—is outrageously optimistic about the mind-refuting implications of the physical sciences); *see also* Stephen J. Morse, *Inevitable Mens Rea*, 27 HARV. J.L. & PUB. POL'Y 51, 55 (2003) (pointing out that "[i]n a thoroughly causal world, if causation were an excuse, no one would be responsible for anything, whether or not we are intentional creatures. But unless proponents of this argument are more successful than incompatibilists in the determinism debate and can furnish convincing reason why causation should excuse, they provide no reason to jettison responsibility practices. Finally, contrary to popular belief, the causation argument is erroneous as an explanation of present responsibility concepts and practices. I have termed this the 'fundamental psycholegal error.' Causation, even by a biologically abnormal structure or process, is not a present excusing condition in western morality and law."). *Cf.* Churchland, *supra* note 1.

[89] *See* Emad H. Atiq, *How Folk Beliefs About Free Will Influence Sentencing: A New Target for the Neuro-Determinist Critics of Criminal Law*, 16 NEW CRIM. L. REV. 449 (2013) (explaining, and railing against, the pernicious influences of folk-psychology, and suggesting several interesting materialist-utilitarian solutions). Note that these modifications would also be possible under a utilitarian system, but at the costs described above.

[90] But see Anders Kaye, *Does Situationist Psychology Have Radical Implications for Criminal Responsibility?*, 59 ALA. L. REV. 611 (2008) for a critique of this assertion. Kaye's position is that

> [r]ecent empirical research shows that human reasoning deviates in important and predictable ways from normative or logical reasoning. One example of this is that we rely upon a variety of heuristics—short cut rules of thumb that substitute for more sophisticated analysis—to help us analyze complex data. Although heuristics frequently lead us to erroneous conclusions about that data analyzed, our use of heuristics is nevertheless generally seen as an 'adaptive' sort of 'bounded rationality,' insofar as it enables us to make quick, moderately reliable judgments about important events when careful reasoning would be too time-consuming, costly or impractical. To

human beings, in their mechanical mystery, have always put less weight in what is true than what is right. On such a scale, utilitarianism cannot outweigh retributivism for sheer usefulness, and hence retributive punishment is more useful even if false than utilitarian punishment is even if true.

IV. CONCLUDING THE PHILOSOPHICAL DISCUSSION

This new model could be intellectually, or practically, tested in several ways. For example, a model statute, in the style of the American Legal Institute's Model Penal Code, setting out a defense like Professor Delgado's proposed (but only sketched) defense of "extreme environmental deprivation"[91] or Judge Bazelon's "Rotten Social Background."[92] Along the same lines, proposed jury instructions or evidentiary rules based on neuroscience would also allow a basic, yet specific and technical outlet for open and bold discussion of these issues in practice. Above all, however, we must proceed with Professor Delgado's exhortation in mind, and "write for those willing to take the risk of appearing thoughtful in these matters."[93]

Therefore, let us attempt to trace a model Rotten Social Background defense statute below.

V. A MODEL STATUTE

As noted *supra*, a re-justification of retributive punishment on utilitarian grounds would allow greater inclusivity for novel defenses, as well as deeper investigation into the social, environmental, and genetic factors influencing criminal behavior.

This note proposes a model statute in the style of the American Legal Institute's Model Penal Code for the Rotten Social Background Defense. Though the statute itself will be dissected and analyzed more thoroughly below, it bears noting here that this model statute embodies only one of several possible approaches to Rotten Social Background.

* * *

> say that heuristic thinking is adaptive, however, is not the same as saying that it fits easily with our working conceptions of the normal human actor.

Id. at 631.

[91] *See* Delgado, *Rotten*, *supra* note 22 (exploring the possibility of a sociologically-derived novel defense based on severe deprivation and neglect in childhood and adolescence).

[92] United States v. Alexander, 471 F.2d 923, 957-65 (D.C. Cir. 1973) (Bazelon, C.J., dissenting).

[93] Delgado, *Wretched*, *supra* note 24, at 15, (quoting Jeremy Waldron, 2009 Oliver Wendell Holmes Lectures, *Dignity and Defamation: The Visibility of Hate*, 123 HARV. L. REV. 1596, 1600 (2010)).

Defense of Extreme Environmental Interference with the Ordinary Functions of Behavioral Control – A criminal defendant presents a prima facie defense of extreme environmental interference with the ordinary functions of behavioral control when he or she shows, by a preponderance of the evidence, that:

(A) He or she lacked responsibility for his or her criminal conduct at the time of such criminal conduct, where he or she shows, by a preponderance of the evidence:

(1) a traceable connection between the criminal conduct and the loss of control causing the criminal conduct;

(2) a traceable connection between the loss of control and the extreme environmental interference causing the loss of control; and,

(3) a close and substantial relationship between the nature and qualities of the extreme environmental interference causing the loss of control and the nature and qualities of the external stimulus which was the proximate cause of the criminal conduct.

(B) Upon presentation of a prima facie case of extreme environmental interference with the ordinary functions of behavioral control, the court shall consider the following factors:

(1) whether the extreme environmental interference is traceable to acts by the defendant which are not traceable to extreme environmental interference;

(2) whether the extreme environmental interference is the result of acts by third parties which amount to coercion to adopt behavioral norms which conflict with prevailing societal norms;

(3) whether the extreme environmental interference is the result of environmental conditions which bear a close and substantial relationship to decisions by society as a whole to extend or deny benefits and burdens in a disparate way to the community or class of which the defendant is a part; and,

(4) any other factors which fairness or justice require.

(C) Upon consideration of the factors enumerated in subsection (B), the court shall instruct the jury:

(1) to consider the factors enumerated in subsection (B) in determining whether the criminal defendant committed a voluntary act beyond a reasonable doubt when he engaged in the criminal conduct with which he is charged; and,

(2) to consider whether, upon the evidence presented, the defendant can be fairly responsible for:

(a) less than 25% of the blame, in which case the jury must render an acquittal unless in a capital case;

(b) less than 50% of the blame, in which case the jury may render an acquittal or return a verdict for a lesser offense at its discretion unless in a capital case; or,

(c) less than 90% of the blame, in which case the jury may return a verdict for a lesser offense at its discretion unless in a capital case.

* * *

VI. ANALYZING AND JUSTIFYING THE MODEL STATUTE

Part V walks the reader through the model statute proposed above. Each section and subsection of the statute is explained, tied to discredited classical theories, contemporary neuroscientific discoveries, or the philosophical controversies surrounding punishment in a materialistic world, as appropriate. Part V will proceed section by section.

A. Title: "Defense of Extreme Environmental Interference with the Ordinary Functions of Behavioral Control"

This title is intended to illustrate the difference between mere "background evidence" of the kind rejected by the United States Supreme Court in *Buchanan v. Angelone* (discussed *supra*), among other cases, and the kind of background evidence which goes directly to a determination of criminal agency.[94]

Professor Delgado first attempted to distinguish his proposed defense based on "extreme environmental deprivation" from the colloquialistic "Rotten Social Background" defense *articulated* by Judge Bazelon in *United States v. Alexander* but actually named and first raised by Murdock's defense counsel at trial in that case.[95]

This title goes farther, and is even more explicit in pointing out that the defense is not merely that the defendant *has* a rotten social background, but that the defendant's rotten social background was so invasive into the ordinary development and functioning of his brain that it deprived him of the ability to control and regulate his behavioral responses to external stimuli in an appropriate, and, most importantly, a non-criminal way.[96] The title is syntactically designed to suggest that the *active* role is played by the environmental interference, which is, one will note, "extreme."[97] The defendant himself is the object of that extreme environmental

---

[94] *Compare* Buchanan v. Angelone, 522 U.S. 269, 273-74 (reasoning that defendant's "uncontrollable" impulse to drink alcohol, a habit that often resulted in bad behavior, was not sufficiently "uncontrollable" to merit irresponsibility mitigation), *with Alexander*, 471 F.2d at 957 (Bazelon, C.J., dissenting) (tracing a troubling connection between some specific events in the defendant's early life and the specific events—the racial slurs—that led to the murder). *But see* Jayaraman, *supra* note 25, at 330-31 (considering the evidentiary problematic of proving specific prior events leading to the instant bad act).

[95] *See supra* Part II.B; *Alexander* 471 F.2d at 960.

[96] The term "Rotten Social Background," though wonderfully evocative, is—for that very reason—undesirable. First, it has a sinister undertone of racism or classism, suggesting that the entire stratum from which the defendant comes—his "society" of outcasts and outsiders—is rotten in a way that "ordinary" society is not. Second, though related, it is thematically under inclusive in that it suggests that the only defendants who would be entitled to the benefits of the defense would come from rotten societal groups. There is no reason to exclude the economically rich from protection, as they may be just as buffeted by maltreatment, neglect, and coercive role-restriction as the so-called underclasses. Delgado's term for the same defense is likewise wanting.

[97] I submit that the title provided—extreme environmental interference with the ordinary functions of behavior control—describes exactly what is going on, rather than illustrating it prettily (if the reader will pardon the expression). The extreme environmental deprivation needs to be much, much more than that alone. To be a plausible novel defense it must justify itself by insisting that the novel aspect (the defense based on environmental causation) be intimately, clearly, closely linked to the classical, well-grounded part (the defense of irresponsibility). All of this goes to the rejustification of retributivism discussed above.

interference, rather than a willing co-participant. This phraseology sets the stage for the rest of the statute, and begins the integration of retributive punishment into the neuroscientifically-supported materialistic worldview that, as explained above, seriously undermines the traditional justifications underpinning classical retributive justice.[98]

B. Opening Clause

> "A criminal defendant presents a prima facie defense of extreme environmental interference with the ordinary functions of behavioral control when he or she shows, by a preponderance of the evidence, that:"

The opening clause of the statute is intended to set out the preliminary evidentiary burden the defendant must meet. There are several important points to consider.

First, when the defendant satisfies section A (discussed *infra*), he presents a prima facie case. Thus, having done so successfully, he will operate under sections B and C (also discussed *infra*) of the statute *unless* the prosecution is able to rebut the prima facie case by showing that the defendant has failed to meet his burden to show all of the elements set forth in subsections 1 through 3 of section A. This allows the prosecution to produce evidence demonstrating, for instance, that the relationship between the defendant's conduct and the extreme environmental deprivation allegedly suffered is too attenuated, or not shown by a preponderance of the evidence, to satisfy section A. However, this would only be possible, or necessary, *after* the defendant had presented her prima facie case. Burden shifting statutes of this kind are both common and effective.[99]

Second, as touched on in the preceding paragraph, the defendant must make the requisite showing under section A by a *preponderance* of the evidence. This is a common technique in criminal proceedings,[100] as generally the only occasion when the standard is "beyond a reasonable doubt" is on the question of actual guilt.[101] A preponderance standard, I submit, creates a fair and equitable balance between defendant and prosecution.

A defendant who can show, with respect to each subsection under section A, that he suffered extreme environmental interference with the ordinary functions of behavioral control, should be permitted to proceed to section B (discussed *infra*), whereas a defendant who is unable to make the requisite showing, again with respect to *every* element of section A, should be denied the right to make a Rotten Social Background claim, as, if it is not more likely than not that he suffered such extensive and extreme environmental deprivation as to be deprived of the ordinary and normal ability to control his behavioral responses to environmental stimuli, his situation is not as strongly related to the re-justifications of retributive punishment discussed

---

[98] *See supra* Part II.C.

[99] *See* McDonnell Douglas Corp. v. Green, 411 U.S. 792, 802 (1973) (setting up a burden shifting test for employment discrimination actions); State v. Burnside, 797 N.E.2d 71, 75 (Ohio 2003) (setting up a burden shifting test for the admissibility of BAC test results).

[100] Lockhart v. Nelson, 488 U.S. 33, 35 (1988).

[101] *Id.*

above.[102] That is, if he is unable to show by a preponderance of the evidence that he had limited or circumscribed responsibility, then he is not really deserving of a defense justified in part by an expansion of retributive ideas allowed by an increased focus on utility and efficiency.[103] If he has not been deprived, there is no need to consider the utility of attributing to him a level of mental responsibility, which may not, in fact, be commensurate with his actual situation.

Determination of whether or not the defendant has met the requisite burden is probably most effective as within the discretion of the trial court, and reviewable as a matter of fact.[104] Sufficiently convincing trial counsel might be able to summon the empathy and pity of a jury more effectively than she could sway an impartial judge whose role is only to ensure that each and every element of section A has been met by a preponderance of the evidence. Additionally, jury-determination of the question would unduly extend the length of trial and the complexity of the decisions left to the jury.

C. Section A

The purpose of this section is to restrict the kinds of extreme environmental deprivation that activate the defense exclusively to those that have a close relation to the crime charged. Thus, beginning with the opening clause of section A, the statute requires the defendant to show that he lacked responsibility *at the time of the criminal conduct*.[105] This begins the thematic framework of the entire section because it requires the defendant to connect the extreme environmental interference he alleges he suffered to the actual crime he in fact later committed.[106]

Subsection 1 continues this trend. Subsection 1 requires the defendant to show a traceable connection between the criminal conduct of which he has been charged with committing and the loss of control or responsibility he claims to have suffered *at the time* (as required by the opening clause of section A). It is submitted that requiring only a traceable connection eliminates some of the evidentiary concerns raised by other scholars and commentators dealing with the problematics of the Rotten Social Background defense.[107] The inclusive "and" in subsection 2, however, raises the burden again, thereby making up for the previous lowering of the burden of showing a connection between conduct and loss or lack of control.

Subsection 2 then requires the defendant to show an additional traceable connection between the loss of control (which was shown in subsection 1 to itself have a traceable connection to the criminal act committed) and the extreme environmental interference alleged to have been suffered. As with subsection 1, requiring the connection to be traceable, as opposed to "close and substantial" or

---

[102] *See supra* Part II.C.

[103] *Id.*

[104] Thus, like evidentiary rulings and other non-dispositive trial-level decisions, the trial judge, who is present to view the demeanor of the witnesses and the peculiar and specific quality of the evidence, is entitled to some deference from a reviewing court, but not when that discretion is abused.

[105] *See* Delgado, *Myth*, *supra* note 25, at 896; Delgado, *Wretched*, *supra* note 24, at 14.

[106] *See infra* Part V.D.-E.

[107] *E.g.*, Robinson, *supra* note 25; Luna, *supra* note 25; Jayaraman, *supra* note 25.

"directly related" or even "proximately caused" makes up for the evidentiary burden of proving that the defendant, who may well be a middle-aged man (as the statute contemplates no restrictions on age or temporal distance), suffered deprivations in a childhood three or four decades past.[108]

Yet, as noted above, the inclusive "and" and the end of subsection 2 makes it possible for a prime facie case to be presented only when the defendant has met all three separate showings by a preponderance of the evidence. Thus, while meeting any given one of the elements may be relatively easy (especially when compared to the difficulties of meeting the elements of an insanity defense, or even of an Model Penal Code extreme emotional disturbance mitigation defense) the burden of meeting all three at once is relatively high. Therefore, most defendants will be excluded from protection under the defense.[109] Only those who have been so traumatized as to be able to separately meet all of the elements of section A will be permitted to move on to section B.

Finally, subsection 3 completes the requisite evidentiary burden for a prima facie case. Subsection 3 is the only subsection that requires the defendant to show a "close and substantial" relationship. This is because it is beneficial to restrict the defense not only to defendants who are able to meet the traceable connection burden between the criminal act committed and the environmental deprivations suffered, but also the defendants who can clearly show that the external stimulus presented, to which the criminal act committed was an uncontrollable response, is related in a clear and direct way to the kinds of deprivations suffered originally.[110]

A pair of examples will help to illustrate the importance of this point. Imagine two non-white, low-income defendants, both of whom have been charged with assault with a deadly weapon, and both of whom have satisfied subsections 1 and 2 of section A. Both defendants were neglected by their parents in early childhood, both were exposed to domestic violence, drug abuse, violent criminal behavior, and constant deception and manipulation. Both were constantly harassed by the police, and were involved in racially motivated attacks on them personally. One has been charged, almost like Murdock in *United States v. Alexander*, with shooting a white police officer who performed a questionable stop-and-frisk that culminated in a fistfight and a brief exchange of gunfire when the officer found a small pistol on the defendant's person.

Such a defendant could potentially meet the burden required by subsection 3 of section A. The second defendant robbed an elderly neighbor at gunpoint. He could

---

[108] As some commentators have noted, the complexities of the inner working of the mind will in many cases defy comprehensive understanding and yet, contemporaneously, will permit a huge number of probable sources for any given set of behaviors, whether desirable or undesirable.

[109] That is, most defendants will not be able to show either that the environmental influence was sufficiently extreme, given the evidentiary problems involved, or will not be able to show that the connection between the extreme environmental interference and the later behavior is sufficiently close to merit protection under the defense. It bears mentioning again that this is less a defense for the extremely poor and more a defense for the extremely maladjusted.

[110] Thus mitigating one of Judge Bazelon's original concerns in *Alexander*. *See* United States v. Alexander, 471 F.2d 923, 962-65 (D.C. Cir. 1973) (Bazelon, C.J., dissenting) (worrying that any expansion in this context of the irresponsibility defenses would either have to be arbitrarily restricted or could potentially encompass almost every criminal actor).

probably not meet the burden in subsection 3. The purpose of the statute is not to create an excuse for every impulsive criminal act committed by every person who might have been mistreated as a child, but rather to allow some defendants who, when confronted with stimuli very similar to those with which they had been tormented throughout early life, act out of uncontrollable fear and anger driven by the amygdala[111] (discussed *supra*) and without sufficient (or any) control from the rational, behavior-modulating prefrontal cortex.[112]

D. Section B

The purpose of this section is to provide a *judicial* basis for evaluating the merits of the prima facie case before the jury is instructed on the proper factual considerations (as set forth in section C). Thus, the defendant, at this stage, will have presented a prima facie case and the state will have attempted (but, presumably, failed) to rebut one or all of the elements of section A. The merits of this framework merit some discussion.

Subsection 1 essentially mandates that the court consider whether the evidence suggests that the extreme environmental interference is traceable to the defendant's own behavior. This is analogous to the Model Penal Code's arrangement for duress, where the defense is unavailable to a defendant who negligently placed himself in a situation that foreseeably led to his being subjected to duress. Thus, subsection 1 prevents easy manipulation of the defense, as a defendant who has foreseeably caused his own extreme environmental interference (i.e. the element or stimuli from the environment that is disrupting the defendant's ability to make meaningful, socially acceptable choices is the defendant *himself*) will not be protected.

This is reconcilable with a philosophical position based on and supported by materialist neuroscience as long as the re-justification of retributive punishment set forth above is assumed present. This section, under that view, distinguishes between unverifiable free will attributions that are not socially beneficial (i.e. that have little utility) and those that do. Thus, a defendant who has foreseeably, even if deterministically, undermined his own position is not protected, while one who has unintentionally, even if deterministically, been undermined, is protected by the statute.[113]

Section 2 goes in the opposite direction. This section is intended to permit judicial consideration of outside interference *not* coming from the defendant's own acts which, if present, would make irresponsibility *more* likely.[114] This is analogous

---

[111] *See supra* note 52.

[112] *See* CARTER, MAPPING, *supra* note 29, at 93-97.

[113] This, again, is one of the many areas of practical consideration made possible only through rejustifying retributivism on utilitarian grounds. When free will attributions advance the good of society as a whole, they may be made even if empirically questionable.

[114] *See* Michael Corrado, *Addiction and Causation*, 37 SAN DIEGO L. REV. 913, 942 (2000).

> There are two sorts of thing to be said, then, about the subjective irrationality explanation of addiction. The first is that if the addict did systematically draw the wrong practical conclusions from his beliefs and desires, he might be entitled to an excuse. But it would not be because of the difficulty of doing otherwise, which after all is what we are trying to capture here. It would be because of a kind of insanity or incompetence. The second response is that that condition does not fit very well with our picture of addiction; the addict is not a bumbling, helpless wanderer who can't fit

to the American Legal Institute's Model Penal Code requirement for duress—the criminal act done by the defendant must have been the result of pressure from a third party.

Section 3 looks to society as whole, and seeks to determine whether it is the inherent imbalances in resource availability, job opportunities, education, and so on, that have created the environmental interference.[115] To the extent that this is so, the judicial decision should weigh more heavily toward a finding of irresponsibility.

Section 4 brings any other noteworthy factors about the specific defendant within the purview of the defense. This section permits a case-by-case appraisal of the degree to which the environment at issue interfered with the defendant at issue. By opening up this avenue of inquiry, the judiciary will be made freer to *disregard* the borderlands of the doctrine at the discretion of the individual judge. Because all of these considerations are evidentiary rulings, they will be reviewed for an abuse of discretion. Ideally, this will allow more willing participation in the statutory scheme, as judges will not feel pressured unduly to be lenient. This is especially important where state judges are elected, and may have to deal with pressures from the electorate to be "tough on crime."

E. Section C

Section C is presented as the most humble preliminary suggestion for attributing and apportioning blame. The proffered numbers are intended more to suggest a baseline matrix from which more precise numbers can be developed. Its inspiration is the apportionment of blame often carried out in comparative-negligence states.

While the specific numbers are negotiable in fact, in theory their purpose is to permit a jury-decision about the degree to which society, as represented by the jurors, should be held accountable for the actions—to whatever extent decided—of the accused.

VII. CONCLUDING REMARKS

The philosophical transformation above is presented as a possible route out of the quagmire that has trapped the RSB discussion. Similarly, the statute proposed above is not intended as a drafting guide for state or federal legislators, but rather a secondary, additional route for leaving the ultra-theoretical realm heretofore the location of RSB discussion, and moving to a more concrete, particularized locale. This, after all, is the most important thing, for we cannot any longer simply brush aside RSB, or other, like defenses. We must be clear on why we reject them—if we do—clear on why we accept them—if we do—and clear on what we want, and what we can get, from a vision of the criminal law in a world where neuroscience has circumscribed free will. The stakes are higher than only brief thought on the issue would suggest. For behind the philosophical dispute lies an actual dispute between society and its individual members, between the State and its citizens, and between the two poles of human feeling, revenge and forgiveness. If we succeed, we can begin to ensure that all humans are given a fair shake. If we fail, we will set up a

---

> his actions to his desires. Addictive behavior, however destructive, is purposeful and often efficient in satisfying the addict's immediate desires.

Id.

[115] Delgado, *Wretched*, *supra* note 24, at 11, 14.

situation in which the people, knowing the State will not ever forgive them, will begin never to forgive it. This is a final situation to be avoided at any cost. The way to avoid it is to discuss the issues *underlying* the actual practices; that is, we must make decisions about philosophy before we can make decisions about policy.